

Métodos de segmentación y seguimiento de estructuras para vídeo laparoscópico

P. Sánchez-González^{1,2}, M. Luna Serrano^{1,2}, A. Fernández Pérez^{1,2}, I. Oropesa García^{1,2},
J.A. Sánchez Margallo³, F.M. Sánchez Margallo³, E.J. Gómez Aguilera^{1,2}

¹ Grupo de Bioingeniería y Telemedicina (GBT), ETSI Telecomunicación, Universidad Politécnica de Madrid, España, {psanchez,mluna,afernandez,ioropesa,egomez}@gbt.tfo.upm.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

³ Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, {jasanchez,msanchez}@ccmijesususon.com

Resumen

En este artículo se presenta el diseño, implementación y evaluación de tres métodos que permiten realizar la detección y el seguimiento de estructuras de interés seleccionadas por el usuario a lo largo de un conjunto de fotogramas de vídeo quirúrgico. El objetivo de estos métodos es la extracción de la información relativa a las estructuras presentes en una determinada escena quirúrgica en entornos de formación o durante los procedimientos de mínima invasión. Los resultados muestran su directa aplicabilidad a entornos didácticos, por ser técnicas semiautomáticas en las que se requiere interacción del usuario.

1. Introducción

La introducción en rutina clínica de la cirugía de mínima invasión (CMI) ha llevado consigo la necesidad de adquisición de nuevas habilidades quirúrgicas, y con ella, el desarrollo de nuevos métodos, algoritmos y dispositivos de formación [1] [2]. Por otra parte, los sistemas de navegación y cirugía guiada por imagen proporcionan al cirujano información necesaria sobre las estructuras que se deben intervenir o preservar y los márgenes de seguridad que se deben respetar durante las intervenciones. Una novedosa investigación aboga por potenciar el empleo del vídeo quirúrgico tanto fuera como dentro del quirófano, precisándose el desarrollo de métodos y algoritmos para analizar y procesar las imágenes que son capturadas por el endoscopio [3].

Un objetivo de este análisis de imágenes de laparoscopia es alcanzar la detección y el seguimiento de los objetos de interés de la escena quirúrgica (fundamentalmente instrumental y estructuras anatómicas) [4][5]. Este análisis puede realizarse con dos objetivos principales: realizar el seguimiento de un determinado órgano y del instrumental durante una intervención quirúrgica o facilitar la docencia de estas técnicas mediante el análisis de los movimientos del instrumental, tanto en entornos de formación (simuladores físicos) como de quirófano [3].

El diseño, desarrollo e implementación de métodos de segmentación de estructuras anatómicas es un campo de alto interés para la investigación en CMI. En función del tiempo de ejecución y otros parámetros que evalúan la precisión y la exactitud de la segmentación, estos métodos podrán ser utilizados para el tratamiento de las imágenes

en tiempo real (cirugía guiada por imagen) o para el tratamiento de imágenes con fines educativos.

En este trabajo se han diseñado, implementado y evaluado tres algoritmos de segmentación semiautomáticos de estructuras anatómicas. Estos algoritmos requieren de la interacción con el usuario para realizar una estimación inicial del área a segmentar, por lo que se plantea su uso fundamental para aplicaciones que no requieran de procesamiento automático en tiempo real. En ese sentido, los algoritmos se incluyen en entornos de formación de nuevos cirujanos, integrándolos en herramientas de autoría que permitan la edición de los vídeos quirúrgicos con fines didácticos [6].

2. Metodología

En el presente trabajo de investigación se han diseñado y desarrollado tres métodos semiautomáticos que permiten segmentar y seguir la estructura anatómica seleccionada por el usuario. Para la implementación de estos métodos se ha empleado la librería OpenCV que utiliza el lenguaje de programación C++.

La validación se ha llevado a cabo en un ordenador Intel Core2 Duo a 3GHz, 4 GB de RAM y Windows 7 Professional 32 bit. Para su realización, se ha optado por una metodología basada en una segmentación manual de la estructura anatómica de interés sobre todos los fotogramas. Con el fin de obtener resultados comparables, se emplea la misma base de vídeos quirúrgicos para evaluar todos los métodos desarrollados. Se han utilizado 5 secuencias de vídeo obtenidas por laparotomía sobre un modelo porcino, que recojan gran parte de la variabilidad de las condiciones de las imágenes quirúrgicas: diferentes entornos con zoom, iluminaciones variables, diferentes estructuras anatómicas, etc. Todos tienen una duración aproximada de 1,5 segundos y constan de 36 fotogramas.

Los parámetros que se van a utilizar para evaluar los métodos implementados, descritos en [9] y [10], incluyen: tiempo de ejecución (s), sensibilidad (S), especificidad (E), valor predictivo positivo (VPP) y valor predictivo negativo (VPN).

Para realizar la evaluación se van a utilizar 5 secuencias de vídeo que se describen a continuación:

- Secuencia 1: se visualiza el hígado, la vesícula biliar, parte del intestino y la pinza laparoscópica. El objetivo de la segmentación en este video es la vesícula biliar.
- Secuencia 2: se visualizan varios lóbulos del hígado y parte del intestino. Se va a segmentar uno de los lóbulos del hígado.
- Secuencia 3: se visualiza el hígado, la vesícula biliar y parte de la pared abdominal. La vesícula biliar es el objetivo de realizar la segmentación en el video.
- Secuencia 4: se visualizan dos lóbulos del hígado, bazo y el intestino delgado. Se va a segmentar uno de los lóbulos del hígado.
- Secuencia 5: se visualiza el hígado, el bazo, el intestino, la pared abdominal y la pinza laparoscópica. El objetivo de la segmentación en este video es el hígado.

A continuación se describen cada uno de los tres métodos que han sido diseñados e implementados.

2.1. Método basado en el ratio del histograma.

Este método, histo-ratio, explota la información extraída del histograma y el cálculo de un ratio del mismo con la finalidad de ajustar todo lo posible la región de la imagen donde se va a realizar la búsqueda de la estructura seleccionada por el usuario.

En primer lugar, se realiza un suavizado de la imagen para eliminar el ruido originado por los medios de adquisición de la imagen. A continuación se calcula el histograma de la imagen completa ($bin_hist_imagen_completa$) y el histograma de la región de interés (bin_hist_roi). El ratio del histograma se calcula según lo marcado por la siguiente expresión:

$$ratio = \frac{bin_hist_imagen_completa}{bin_hist_roi}$$

El ratio del histograma permite eliminar de la imagen los colores de la estructura presentes en el resto de la escena quirúrgica. Finalmente, esta información se utiliza para aplicar el algoritmo CAMSHIFT (*Continuously Adaptive Mean Shift*) [7], que calcula la región más probable del nuevo fotograma en la que se puede encontrar la estructura anatómica seleccionada por el usuario.

2.2. Método basado en los espacios de la imagen.

Este método se fundamenta en la recopilación de la distinta información que se obtiene en cada uno de los espacios de imagen. En concreto, se obtiene información de los espacios RGB, GRAY y HSV. A continuación, se obtiene tanto la media (*mean*) como la desviación típica (*stddev*) del valor de la intensidad en la región de interés y se calcula el umbral máximo (*max*) y mínimo (*min*) que determinan las regiones de los siguientes fotogramas que pueden resultar de interés:

$$\min = mean - 1.5 * stddev$$

$$\max = mean + 1.5 * stddev$$

En el caso de que exista una región en el siguiente fotograma cuyos parámetros se encuentran dentro de esos umbrales, se tomará como región de interés y se resaltará en la imagen con un rectángulo o con el contorno de esa región en función de la elección del usuario.

2.3. Método basado en la detección de la forma.

El diseño de este método de segmentación de estructuras se basa en el algoritmo "Template-Matching" [8]. En primer lugar, se extrae de la imagen la semilla seleccionada por el usuario. Esta semilla se utilizará como patrón, por lo que será utilizada para buscar en los sucesivos fotogramas información similar a la contenida en ella. Esta semilla define en cada fotograma una región de interés. La posición de esa región dentro de la imagen se calcula en cada fotograma y se obtiene un conjunto de puntos denominados de interés. Estos puntos indican la localización en la imagen de la estructura anatómica seleccionada. El modelo de visualización empleado incluye el marcado del área detectada con un rectángulo.

3. Resultados y Discusión

En este apartado se va a realizar una descripción de los resultados obtenidos al comparar la segmentación obtenida por cada uno de los métodos con una segmentación manual, empleada como *gold estándar*. A continuación se discute cada uno ellos y se analiza su adecuación a los entornos propuestos.

En la Figura 1 se observa el resultado de la segmentación manual en uno de los fotogramas de los 5 videos utilizados.

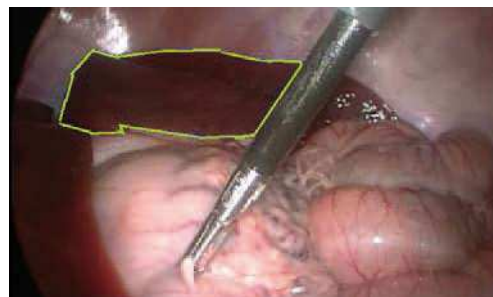


Figura 1. Resultado de realizar la segmentación manual.

3.1. Método histo-ratio.

En la Tabla 1 se puede observar la evolución del valor de cada uno de los parámetros de evaluación en función de la secuencia de vídeo.

Videos	S	E	VPP	VPN
1	21,1%	99,9%	66,5%	99,8%
2	45,0%	90,4%	36,8%	99,7%
3	30,7%	99,9%	46,1%	99,8%
4	20,2%	95,4%	32,3%	74,5%
5	91,5%	99,2%	91,1%	99,2%
Promedio	42,2%	97,0%	48,6%	94,7%

Tabla 1 Sensibilidad, especificidad, VPP y VPN del método histo-ratio.

El tiempo de ejecución de este método ha sido de 10,32 segundos en promedio. La sensibilidad presenta unos

valores constantes y próximos al 45%, con un valor medio de 42,2%. El valor predictivo positivo tiene un valor medio del 48,6% fluctuando los valores entre el 33% y el 91%. Se puede afirmar que a pesar de que los valores no superen el 50%, la estructura segmentada en todos los casos se aproxima correctamente a la estructura de interés en todos los casos. El motivo de que en algunos casos estos valores sean relativamente bajos se debe principalmente a que la superficie de la segmentación no cubre la totalidad de la estructura de interés.

La especificidad media es del 97%, con lo que este método descarta correctamente aquellos píxeles que no corresponden a la estructura anatómica de interés. En cuanto al valor predictivo negativo, el método alcanza unos valores en torno al 100%.

El resultado visual en uno de los vídeos queda recogido en la Figura 2. Al compararlo con la Figura 1, se puede comprobar cómo el área segmentada por el método se ajusta con bastante precisión a la zona segmentada de manera manual.

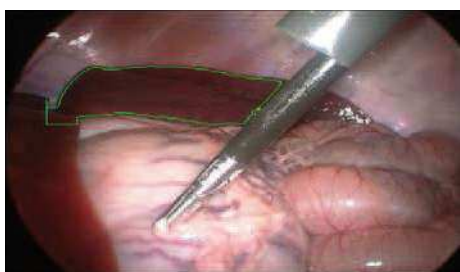


Figura 2. Resultado al aplicar el método histo-ratio.

3.2. Método espacios imagen

La Tabla 2 recoge los resultados obtenidos para cada uno de los vídeos empleados en la validación de los métodos.

Videos	S	E	VPP	VPN
1	10%	99,7%	20%	99,8%
2	12%	90,7%	20%	99,5%
3	17%	99,5%	13,5%	99,2%
4	22%	67,0%	27%	63,7%
5	68%	98,4%	80,3%	97,2%
Promedio	24%	93,0%	32,0%	91,7%

Tabla 2 Sensibilidad, especificidad, VPP y VPN del método espacios-imagen

El tiempo medio de ejecución de este método para las secuencias de vídeo empleadas es de 1,07 segundos. La sensibilidad oscila entre los valores del 10% al 68%, tomando un valor medio del 24%. El valor predictivo positivo fluctúa entre el 20% y el 81%, tomando un valor medio del 32%. La oscilación de estos valores se debe a la estructura anatómica que va a ser segmentada.

La especificidad tiene unos valores muy altos, entre el 67% y el 100%. Su valor medio es de 93%. El valor predictivo negativo tiene un valor medio del 91,7% y fluctúa entre el 64% y el 100%. Los datos obtenidos de estos dos últimos parámetros muestran la eficacia del método, que descarta correctamente los píxeles que no corresponden a la estructura anatómica de interés.

En la Figura 3 se pueden observar los resultados obtenidos en una de las secuencias de vídeo. Para ello se resalta la estructura detectada con un rectángulo:

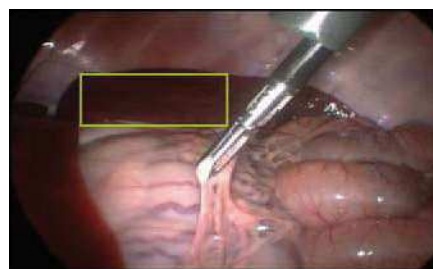


Figura 3. Resultado al aplicar el método espacios imagen.

3.3. Método basado en la detección de la forma.

En la Tabla 3 se puede observar la evolución de los parámetros de evaluación en función de la secuencia de vídeo.

Videos	S	E	VPP	VPN
1	31%	99,7%	70%	99,8%
2	23%	99,2%	59%	99,7%
3	35%	99,7%	28,7%	99,8%
4	17%	99,2%	50%	74,5%
5	91%	99,2%	91,3%	99,2%
Promedio	39%	99,5%	59,4%	94,6%

Tabla 3 Sensibilidad, especificidad, VPP y VPN del método basado en la forma

El tiempo medio de ejecución de este método es de 5,34 segundos. La sensibilidad toma valores entre el 17% y el 91%, teniendo un valor medio del 39%. El valor predictivo positivo medio es del 59,4%. A pesar de la alta variabilidad de estos parámetros, se observa por tanto que la segmentación cubre el área de la estructura anatómica de interés.

La especificidad presenta valores estables, con un valor medio del 99,5% y el valor predictivo negativo medio es del 94,6%. Con estos valores se puede afirmar que el algoritmo descarta correctamente aquellos píxeles que no corresponden a la región de interés en las secuencias de vídeo analizadas.

A modo de ejemplo, la Figura 4 muestra los resultados obtenidos en una de las secuencias de vídeo:

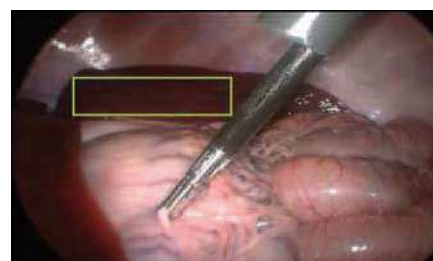


Figura 4. Resultado al aplicar el método basado en la detección de la forma.

La Tabla 4 recoge los valores medios de cada uno de los parámetros de evaluación en los tres métodos implementados. Como se observa los valores de estos parámetros varían en función de la estructura de interés

segmentada. Los mejores valores se han obtenido en las secuencias en las que se ha segmentado el hígado. Esto es debido a que el hígado presenta en todas las secuencias unas características concretas que le diferencian significativamente del entorno anatómico en el que se encuentra, no ocurriendo lo mismo en el caso de la vesícula biliar.

Método	S	E	VPP	VPN	Tiempo ejecución
Histo-ratio	42,2%	97,0%	48,6%	94,7%	10,32 s
Espacios - imagen	24%	93,0%	32,0%	91,7%	1,07 s
Basado en la forma	39%	99,5%	59,4%	94,6%	5,34 s

Tabla 4 Parámetros de evaluación de los métodos desarrollados.

El valor de la especificidad y del VPN se aproxima a valores del 100% en los tres métodos por lo que se puede afirmar que descartan correctamente los píxeles que no pertenecen a la estructura anatómica de interés. Por lo tanto, la estructura anatómica de interés seleccionada no afecta a estos dos parámetros.

En los parámetros relativos a la sensibilidad y al VPP se aprecian las mayores diferencias entre secuencias. Esto se debe principalmente a dos motivos: el primero es el tipo de recuadro que se emplea. En el caso de usar el contorno de la estructura de interés los resultados mejoran en comparación con el rectángulo sobre el área de interés. Esto se debe a que el área cubierta por el contorno es mayor que en el caso del rectángulo y por tanto el número de píxeles que se incluyen dentro de la región de interés es mayor.

A pesar de la variabilidad comentada anteriormente, la región segmentada se ajusta correctamente a la estructura anatómica de interés. Como se puede observar es el método del histo-ratio y el método basado en la forma los que obtienen valores más altos, en torno al 90%.

En cuanto al tiempo de ejecución, el método histo-ratio presenta los valores más elevados, mientras que los otros dos métodos tienen un tiempo de ejecución similar. Estos últimos presentan un coste computacional que los hace viables para su integración en entornos de formación quirúrgicos.

Al ser semiautomáticos, todos los métodos se encuentran condicionados por la semilla introducida por el usuario. Esta influencia es mayor en el método de detección de la forma ya que este método toma las características iniciales de la semilla y no lleva a cabo una actualización del patrón de búsqueda a lo largo del algoritmo.

4. Conclusiones

La incorporación de las técnicas de CMI en rutina clínica lleva asociada la necesidad de desarrollar nuevas metodologías de formación de nuevos cirujanos y de sistemas de navegación quirúrgicos que aporten información relevante durante la propia intervención. El análisis de la imagen captada por el endoscopio permite proporcionar al cirujano información de utilidad acerca de

la escena quirúrgica de manera no intrusiva. Así, su procesamiento permite llevar a cabo la localización y seguimiento de los elementos de interés del escenario.

En este trabajo se han diseñado, implementado y evaluado tres métodos semiautomáticos que permiten detectar y seguir la estructura de interés seleccionada por el usuario. Los tres métodos desarrollados ofrecen buenos resultados para entornos de formación quirúrgicos, aunque para su completa validación es necesaria la realización de un estudio más completo que haga uso de un mayor número de secuencias de vídeo quirúrgicas de distinta naturaleza y duración. La elección final de uso de los métodos deberá hacerse en función del compromiso entre tiempo de ejecución, exactitud, precisión y robustez deseada.

Agradecimientos

El trabajo de investigación ha sido parcialmente financiado por el proyecto intramural del CIBER-BBN THEMIS.

Referencias

- [1] Chekan EG., Pappas .N. General principles of minimally invasive surgery. *Basic Science and Clinical Evidence*, 2001, pp. 429-453.
- [2] Cuschieri A. Laparoscopic surgery: current status, issues and future developments. *Surgeon*, 3(3), 2005, pp.125-138.
- [3] Sánchez-González P, Cano AM, Oropesa I, Sánchez-Margallo FM, del Pozo F, Lamata P, Gómez EJ. Laparoscopic Video Analysis for Training and Image Guided Surgery. *Minimally Invasive Therapy & Allied Technologies* (in press), 2011.
- [4] Cano AM, Vara I, Sánchez-González P, Gómez, EJ. Laparoscopic image analysis for automatic tracking of surgical tools. *Proceedings of Computer Assisted Radiology and Surgery (CARS 2008)*, vol.3 suppl. 1, 2008, pp. S279.
- [5] Sánchez-González P, Gayá F, Cano AM, Gómez, EJ. Segmentation and 3D reconstruction approaches for the design of laparoscopic augmented reality environments. *Lecture Notes in Computer Science*, 2008, pp.127-134.
- [6] Sánchez-González P, Fernández A, Oropesa I, Noguera J, Sánchez-Margallo FM, Burgos D, Gómez EJ. Herramienta de autoría de contenidos didácticos multimedia para entorno de formación colaborativo en Cirugía de Mínima Invasión. *Revista de Educación a Distancia (RED)*, vol. 28, 2011, pp. 1-9.
- [7] Bradski GR. Computer vision face tracking as a component of a perceptual user interface. *In Workshop on Applications of Computer Vision*, 1998, pp. 214-219
- [8] Zhao L, Davis LS. Closely coupled object detection and segmentation. *Computer Vision*, vol. 1, 2005, pp. 454-461,
- [9] Pérez Rodríguez R et al. Segmentación y seguimiento de estructuras en imágenes laparoscópicas. *Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB2008)*, 2008, pp.108-111.
- [10] Gil Rodríguez JL. Evaluación de calidad en la segmentación de imágenes. *Reporte Técnico: Reconocimiento de patrones*. ISSN: 2072-6287